The Issue of Trust in Autonomy

Scott N. MacKinnon

Division of Maritime Studies
Chalmers University
Gothenburg, Sweden

International Seminar on Safety and Security of Autonomous Vessels
21 March, 2018
Delft, The Netherlands
Issues in Shipping Industry

• SUSTAINABILITY
  — Decreased Manning
  — Recruitment and Retention
  — Time in Port (JIT)

• ENVIRONMENTAL CONCERNS
  — Slow Steaming
  — Route Optimization

• INCREASED COMPLEXITY
e-Navigation

A strategy developed by the International Maritime Organization to improve shipping safety and efficiency through better organization of data on ships and on shore, and better data exchange and communication between ships and the ship and shore.
Automation *

1. the method of making a machine, a process, or a system work without being directly controlled by a person

2. automatic working of a machine, process, or system by mechanical or electronic devices that take the place of humans

*Merriam-Webster
Automation *

1. the method of making a machine, a process, or a system work without being directly controlled by a person

2. automatic working of a machine, process, or system by mechanical or electronic devices that take the place of humans

• Decision-support (human in-the-loop)
• Decision-making (human out-of-the-loop)

- We have to think of the shipping industry as a complex socio-technical system (which "now" has both centralized and distributed systems)
The more complex a system is the greater the risk that something, somewhere, sometime will fail.

Ironies of Automation (Bainbridge, 1983)
Automation complexity

Automation surprise

Automation irony

… and that failure might come as a big surprise

Ironies of Automation (Bainbridge, 1983)
Automation complexity

Automation surprise

Automation irony

Automation is most reliable in simple tasks.

The more functionalities needed the less reliable automation becomes.

Ironies of Automation (Bainbridge, 1983)
UMS Method of Control

Control level | Description | Operator | UMV |
---|---|---|---|
1 Human operated | Remote control
(All control by human operator) | Turn to new course | |
2 Human directed | Permissive
(UMV asks for permission) | Go ahead / No go / <silent> | |
3 Human delegated | Declarational
(UMV declares intention) | I would like to turn to new course | |
4 Human monitored | Reportive
(UMV reports action) | No go / <silent> | |
5 Autonomous | Autonomous
(UMV does not report) | I will now turn to new course | |

From: SARUMS (SAFETY AND REGULATIONS FOR UNMANNED MARITIME SYSTEMS)
Trust (Interpersonal)

Disposition toward the world and the people in it (Rotter, 1967)

Socially learned expectations that vary based on social order (Barber, 1983)

A willingness to accept vulnerability (Mayer et al, 1995)

The attitude that an "agent" will help achieve an individual’s goals in a situation characterised by uncertainty and vulnerability (Lee and See, 2004)
Trust* (Human Automation)

Performance-based trust: how well an automated system executes a task

Process-based trust: based on the operator’s understanding of the methods an automated system uses to perform tasks

Purpose-based trust: the designer’s intended use for an automated system

*Hoff and Bashir (2015)
Trust* (Human Automation)

Dispositional trust: represents an individual’s overall tendency to trust automation, independent of context or a specific system (culture, gender, personality, AGE)

Situational trust:
- External: depends largely on the type of system, complexity and difficulty of task
- Internal: interaction with self-confidence

Learned trust: drawn from past experiences of current interactions

*Hoff and Bashir (2015)
The relationship among calibration, resolution, and automation capability in defining appropriate trust in automation.

Maritime Unmanned Navigation through Intelligence in Networks

— Research the feasibility of autonomous unmanned ship and human-center automation governance from shore-based facilities during deep sea voyage

— 3 year EU sponsored Framework 7 project

— Simulated 200 meters long dry bulk carrier
So what did these “Operators” do?
...and who will be the "trusting" people...
The "Generations"

- Silent (1925-1941)
- Baby Boomer (1946–1964)
- Generation X (1960-1980)
- Generation Z (2000-+…)
Generation Z

- Always had the internet
- Comfortable with technology
- Non-Technical Skills?

Grandma... it’s not that hard! Go into Settings... select Wi-Fi... just tap it with your finger... oh for goodness sake!
So where might the errors occur?

Cognitive Psychological School
Swiss Cheese Model

- Unsafe Acts
- Preconditions
- Organisational influences
- Supervision
Figure after Wickens (1992)
Joint Cognitive Systems School

Safety—I and Safety—II
The Past and Future of Safety Management
Erik Hollnagel
Work-as-imagined (What stakeholders believe happens or should happen)

![Diagram of Work-as-imagined]

Work-as-done (What actually happens)

![Diagram of Work-as-done]

What are the implications for the design and execution of work procedures?

What are the implications for recovery from an unexpected event?
Figure after Wickens (1992)
Creating Trust with the User

• Design for appropriate trust, not greater trust.

• Show the past performance of the automation. \( (\text{Big Data}) \)

• Show the process and algorithms of the automation by revealing intermediate results in a way that is comprehensible to the operators. \( (\text{Data Transparency}) \)

• Show how automation is related to the users’/system’s goals. \( (\text{Confirmation of Mental Model}) \)

• \textbf{Train} operators regarding its expected reliability, the mechanisms governing its behaviour, and its intended use. \( (\text{Lessons from Aviation and Nuclear Industry}) \)
Thank you!

Questions?